

**Testimony before the Michigan House of Representatives
Education Committee
May 6, 2014**

**Donald E. Heller
Dean, College of Education
Michigan State University**

Representative Lyons and Members of the Committee:

Thank you for the opportunity to speak with you today. I come before you as an educational researcher and am not representing the views of Michigan State University.

I wish to provide a few comments regarding House Bills 5223 and 5224. I appreciate the fine work that the Michigan Council on Educator Effectiveness performed in coming up with a plan for evaluating teachers and school administrators in the state. Their final report was a thoughtful document, and I know that much of the information contained in these two bills is based on the MCEE recommendations.

As the legislature moves forward with these bills, I would like to add some cautionary advice with respect to the use of student assessments in evaluating educators. Teachers play a very important role in promoting student learning, but they are only a single factor among many that determine how much students will learn in a given subject matter over the course of a year.

A recent statement from the American Statistical Association, the nation's leading professional association on the subject of the use of statistics in research and practice, addressed the use of value-added models in teacher assessment.¹ In its statement, the ASA said:

Research on VAMs has been fairly consistent that aspects of educational effectiveness that are measurable and within teacher control represent a small part of the total variation in student test scores or growth; most estimates in the literature attribute between 1% and 14% of the total variability to teachers. The majority of the variation in test scores is attributable to factors outside of the teacher's control such as student and family background, poverty, curriculum, and unmeasured influences.

HB 5223 calls for 25% of a teacher's evaluation in the next three school years to be based on student growth data, and beginning in 2017-18, 40% of the evaluation needs to be based on student growth. Thus, you are proposing to place into law for this coming year a system that the best statistical minds in the country have concluded can predict at best approximately 14% of the contributions a teacher makes to student test scores. And yet 40% or more of a teacher's evaluation will be based on these measures.

¹ American Statistical Association (2014, April 8). "ASA Statement on Using Value-Added Models for Educational Assessment." Alexandria, VA: Author.

Let me try to provide an analogy for you. We all know that we can step on a scale in the morning and see what we weigh. And if we compare that to our weight a year ago, we can measure how much our weight has grown over the course of that year.

But you would need to be cautious in using that measure of weight growth to determine the nutritional value of the food you had eaten over the course of the year. The nutritional value of what you eat is only one element of your weight change. Research has demonstrated that weight change is also affected by how much you exercise, your genetic factors, and even the amount of stress you endure. And these different factors will have varying influences on any individual person. For some of us, it is what we eat that is most influential; for others, exercise is most important.

The same is true for student test scores. Teacher performance is only one component that contributes to score gains, and as the ASA points out, research shows that it is a relatively small component overall.

As you are probably aware, other states that have already implemented teacher assessment systems that include VAM scores as a major component have run into problems. Teachers in Houston recently filed a lawsuit in federal court against the school district, alleging that the assessment system is inaccurate and unfair. One of the plaintiffs, an award-winning history teacher, was rated as “most effective” one year, and then the next, using the same teaching techniques but with a new class, was rated as having no impact on his students’ learning. A similar suit challenging the statewide teacher assessment system in Florida is in the federal courts in that state. VAM scores in New York City and elsewhere have been shown to be unreliable as a tool for teacher assessment.

In summary, I urge this committee to use caution in moving forward with legislation that puts into place a teacher evaluation system that requires the use of statistical tools that have yet to be shown to be a good measure of teacher performance. Doing so, I am afraid, will make it even harder to attract top students to become the future teachers our state so desperately needs, a goal that we all share. If people do not believe they will be fairly evaluated, it is unlikely they will want to become a teacher. And we will all be worse off because of it.